



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification<sup>4</sup> :</b>  C12P 21/00	<b>A1</b>	<b>(11) International Publication Number:</b> WO 87/ 02061  <b>(43) International Publication Date:</b> 9 April 1987 (09.04.87)
<b>(21) International Application Number:</b> PCT/US86/02074 <b>(22) International Filing Date:</b> 2 October 1986 (02.10.86)  <b>(31) Priority Application Number:</b> 783,787 <b>(32) Priority Date:</b> 3 October 1985 (03.10.85) <b>(33) Priority Country:</b> US  <b>(71) Applicant:</b> BIOTECHNOLOGY RESEARCH PARTNERS, LTD. [US/US]; 2450 Bayshore Parkway, Mountain View, CA 94043 (US).  <b>(72) Inventors:</b> PROTTER, Andrew, A. ; 535 Everett, Apartment 513, Palo Alto, CA 94301 (US). KANE, John, P. ; 1490 Marlborough Road, Hillsborough, CA 94010 (US).		<b>(74) Agents:</b> MURASHIGE, Kate, H. et al.; Ciotti & Murashige, 545 Middlefield Road, Suite 200, Menlo Park, CA 94025 (US).  <b>(81) Designated States:</b> AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> NOVEL LIPOPROTEIN-BASED DRUG-DELIVERY SYSTEMS  <b>(57) Abstract</b>  Pharmaceutical compositions in which the drug delivery system permits selective targeting take advantage of the amphipathic helices of apolipoproteins or binding regions of apolipoproteins for LDL receptors or both. In one embodiment, the LDL receptor binding region of an apolipoprotein, in particular apoB or apoE is covalently bound to either an active ingredient per se or to a carrier for the active ingredient. This delivery system specifically targets tissues which bear the LDL receptor. In another embodiment, the amphipathic alpha helices normally associated often with apolipoproteins are conjugated to proteins which target specific tissues such as immunoglobulins or, preferably, the LDL receptor binding protein regions. This fused system is then associated with an active ingredient or an intermediate carrier therefor. The drug delivery systems described provide novel targeting systems for homing active ingredients to specific tissues.		

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	LI	Liechtenstein	SN	Senegal
CH	Switzerland	LK	Sri Lanka	SU	Soviet Union
CM	Cameroon	LU	Luxembourg	TD	Chad
DE	Germany, Federal Republic of	MC	Monaco	TG	Togo
DK	Denmark	MG	Madagascar	US	United States of America
FI	Finland				

5

NOVEL LIPOPROTEIN-BASED DRUG-DELIVERY SYSTEMSTechnical Field

10           The invention relates to preparations for  
effective, targeted drug delivery. In particular, it  
relates to using the receptor-binding activity of  
certain apolipoproteins, the phospholipid-binding  
ability of amphipathic helices contained in  
15   apolipoproteins, or both, for target-specific delivery  
of active substances to vertebrates.

Background Art

Efficient drug delivery demands that the active  
20   ingredient of the pharmaceutical or other composition be  
carried from its point of administration to its site of  
activity or processing without undue dilution in the  
digestive or circulatory system to undermine its  
effect. It is only common sense that unwanted side  
25   effects of a particular drug will be minimized if it is  
immediately directed to its site of desired activity  
rather than allowed to make its own way through the  
subject's corpus and exert effects in undesired  
locations. This is particularly important, for example,  
30   when toxic substances are used to destroy particular  
unwanted cells such as tumor cells. Most such agents  
are equally, or at least similarly, capable of  
destroying healthy tissue; to the extent they are  
allowed to roam undirected they may result in

-2-

considerable damage. Alternatively, many materials are processed or modified by the liver in order to place them in an effective role, and as a general proposition, the desired delivery to a particular target is mediated by preliminary delivery to the liver. Thus, specific targeting agents may direct the active ingredient either to the ultimate site, or to the liver for intermediate handling.

Various mechanisms have been utilized to target the active ingredients in pharmaceuticals. The most common is the use of antibodies raised against a targeted tissue and linked to the active ingredient, either covalently or by association. These so-called immunoconjugates have elicited wide interest in the last decade. On the other hand, relatively little research has been directed to the use of other specific binding moieties which are expected to attach preferentially to the liver or other target cells.

Some attention has, however, been given to the use of lipoproteins in drug delivery. These proteins, classified as chylomicrons (CM), very low density lipoproteins (VLDL), low density lipoproteins (LDL), and high density lipoproteins (HDL) according to their behavior on gradient centrifugation are the major vehicles for transport of lipids in the plasma. The characteristics and compositions of these four plasma fractions have been reviewed in connection with their use in drug delivery by Counsell, R.E., et al, J Med Chem (1982) 25: 545-550. The two lighter fractions, CM and VLDL, appear to be the major carriers for triglycerides; LDL and HDL appear to be delivery systems for cholesterol. All contain proteins on their surfaces, and their composition with respect to the various apolipoproteins thus contained varies according

to the particle. Of particular interest with respect to the invention herein is the LDL fraction, which contains more than 80% apolipoprotein B (apoB) and some apolipoprotein E (apoE). Both of these proteins appear  
5 to be involved in the binding of the LDL to specific receptors in the liver and in the binding of the LDL to receptors on various cultured cell lines.

The lipoproteins in general contain a lipid center surrounded by a monolayer containing  
10 phospholipids and the apolipoproteins. Regardless of the nature of the specific binding of the surface apolipoproteins to receptors, attempts have been made to utilize the lipoproteins as delivery vehicles for drugs by sequestering an active substance in the lipid  
15 fraction. It was demonstrated by Rudman, D., et al, Pharmacol Exp Ther (1972) 180: 797, that various materials of pharmaceutical interest, including estradiol, testosterone, and pentobarbital, can be sequestered in the lipoprotein fraction. Reconstituted  
20 LDL has also been shown to be an effective vehicle for the delivery of fluorescent probes to cells for the purpose of labeling (Krieger, M., et al, J Supramol Struct (1979) 10: 467-478), and reconstituted LDL containing cytotoxic agents has been used to deliver  
25 these agents specifically to tumor cells in vitro (Rudling, M., Canc Res (1983) 43: 4600-4605). Selective in vitro activity against cancer cells using nitrogen mustards sequestered in LDL was also demonstrated by Firestone, R.A., et al, J Med Chem (1984) 27:  
30 1037-1043.

Some specific studies of the nature of the apolipoproteins which make up LDL, apoB and apoE, have been made. ApoE is a 299-residue protein of known amino acid sequence that is a component of several of the

-4-

lipoproteins. The binding domain of apoE has been studied in relation to the receptor sites (Innerarity, T.L., et al, J Biol Chem (1983) 258: 12341-12347; Weisgraber, K.H., et al, ibid, 12348-12354). A number  
5 of isoforms of this protein have been found.

Apolipoprotein B (apoB) is much larger. When deprived of its lipid content, the protein appears to have a molecular weight of about 380,000, corresponding to almost 4000 amino acid residues. This protein  
10 appears to be present in two major forms, the larger peptide designated B-100, which is associated with hepatic lipoproteins, and a somewhat smaller form designated B-48, which is associated with intestinal lipoproteins and the chylomicrons. The hepatic form,  
15 B-100, appears to be processed by proteolytic enzymes in the plasma into two fragments designated B-74 and B-26.

While the entire amino acid sequence of apolipoprotein E is known, and a cDNA clone encoding the entire protein has been isolated (McLean, J.W., et al, J Biol Chem (1984) 259: 6498-6504), the extremely large  
20 size of apoB has prevented similar progress. Even the exact molecular weight is not known with certainty. However, cDNA clones encoding small portions of the protein sequence have been recovered from a human cDNA  
25 library (Deeb, S.S., et al, PNAS (USA) (1985) 82: 4983-4986) and from rat liver (Lusis, A.J., PNAS (USA) (1985) 82: 4597-4601). The Deeb sequence encodes approximately 200 amino acids; the Lusis sequence, only about 80.

30 The receptor for, and receptor-binding properties of, these proteins have been studied to a limited extent. Yamamoto, T., et al (Cell (1984) 39: 27-38) examined the LDL receptor site using a full-length cDNA clone for human LDL receptor cDNA

isolated from a human fetal adrenal cDNA library. On the basis of this clone, it appears that the receptor is an 839 amino acid protein having five identified domains, with the presumed ligand binding site  
 5 consisting of a 322 amino acid cysteine-rich region containing eight negatively charged side chains. These results are consistent with the finding of Weisgraber, K.H., et al (J Biol Chem (1978) 253: 9053-9062) that LDL modified to block the positive charge on the lysyl  
 10 residues is incapable of competing with native LDL for sites on the receptor. This failure to compete is reversible by removal of the blockage of positive charge. Reductive methylation to modify the lysyl side chains while failing to neutralize the positive charge,  
 15 irreversibly destroys the ability of LDL to bind to the receptor. An attempt to map the LDL antigen with monoclonal antibodies has also been reported (Tikkanen, M.J., et al, J Lipid Res (1982) 23: 1032-1038).

In addition to receptor-binding ability  
 20 exhibited by certain apolipoproteins, apolipoproteins in general appear to contain specialized, but generic, structures to enable them effectively to bind phospholipids. These structures are considered to be amphipathic alpha helices which contain oppositely  
 25 facing hydrophobic and hydrophilic regions (Sparrow, J.T., et al, Annals New York Acad Sci (1980) 348: 187-211). Such regions have been found in apoCIII, apoCI, apoCII, apoAI (see also McLachan, A.D., Nature (1977) 267: 465-466), and apoAII and apoE. The presence  
 30 or absence of such domains in apoB has not been discussed.

As described above, the desirability of employing lipid transport systems for targeted drug delivery has been recognized, but no controlled

efficient method has been found. The nonspecific delivery of drugs by lipid-based systems, notably by liposomes, is now a crowded technology, but suffers from the inherent disadvantage of lack of targeting and specificity. The compositions and methods of the present invention provide new approaches to targeted drug delivery using the mediation of lipid transport.

#### Disclosure of the Invention

10           The invention provides drug delivery systems which are both specifically targeted and lipid compatible. The approach is two-fold. In one approach, the receptor recognition domains of apolipoproteins, in particular of apolipoprotein B or E, are used to serve  
15 as a binding moiety, and are covalently linked to additional substances which are either carriers or themselves active ingredients. In an alternate approach,  $\alpha$  helical amphipathic regions such as those of apolipoproteins are used to bind directly to  
20 hydrophobic active ingredients or to bind lipid compatible carriers containing active ingredients. These amphipathic regions are in turn bound to specific targeting substances including the receptor binding regions of apoB or apoE as well as to completely  
25 different binding moieties such as antibodies. A particularly preferred approach combines the receptor binding regions of the apolipoproteins with their amphipathic helices in a single protein which then serves as a directed carrier of a lipid soluble active  
30 ingredient.

Thus, in one aspect, the invention relates to a composition for delivery of an active substance which comprises an LDL receptor binding region such as that of apoB or apoE covalently bound to a substance which is

either itself an active ingredient or a carrier for an active ingredient. In another aspect, the invention is directed to a composition for delivery of an active substance which comprises a lipophilic peptide

5 containing at least one amphipathic  $\alpha$  helical domain conjugated covalently to a binding moiety, i.e., a compound that has specific binding regions, such as an immunoglobulin or an LDL receptor binding region. The compositions and methods of the invention are suitable

10 for use in any vertebrate subject including, and preferably, mammalian and human subjects. Other aspects of the invention include methods of administering active ingredients employing the compositions of the invention and specific novel components of these compositions.

15 The novel components include the LDL receptor binding region of the apoB protein and the amphipathic regions thereof, and fusion proteins comprised of LDL receptor binding region and amphipathic regions derived from alternate apolipoproteins, of amphipathic regions and

20 specific binding moieties, such as antibodies, or of LDL receptor binding region and active substance.

In still other aspects, the invention is directed to materials and methods useful in constructing the drug delivery compositions of the invention,

25 including the relevant DNA sequences, vectors, expression systems, and transformed hosts.

#### Brief Description of the Drawings

Figure 1 shows the amino acid sequence of the

30 N-terminal portion of apoB-26 with putative LDL-receptor binding regions and amphipathic helices indicated.

Figure 2 shows the DNA and deduced amino acid sequence for a cDNA insert encoding the N-terminal region of human apolipoprotein B through amino acid 266.

Figure 3 shows the DNA sequence and deduced amino acid sequence for the N-terminal region of human apolipoprotein B through amino acid 430.

Figure 4a shows the base sequence of DNA used  
5 in the linking region of a fusion protein containing portions of apoB and apoAI; Figure 4b diagrams the portions of apoAI genomic DNA used.

Figure 5 shows the DNA sequence and deduced amino acid sequence of a fusion protein containing apoB  
10 receptor binding and apoAI amphipathic regions.

#### Modes of Carrying Out the Invention

##### A. Definitions

As used herein, "LDL receptor binding region"  
15 refers to a single peptide sequence which is capable of binding to LDL receptors. A protein fits the definition of containing an LDL receptor binding region if it is capable of competing with iodine-125 labelled LDL in binding to surface receptors on human fibroblasts as  
20 described by Goldstein, J.L., et al, Cell (1976) 7: 85-95 or to receptors on isolated liver membranes as described by Basu, S.K., et al, PNAS (1976) 73: 3178-3182.

As stated in the Background section above,  
25 something is now known about the nature of these binding regions. In general, they appear to contain sectors of positive charge, generally characterized as segments of the protein containing multiple lysine (or arginine) residues in relatively exposed portions of the protein.  
30 Segments containing these residues are often bracketed by cysteine residues, thus providing substrates for disulfide linkages to consolidate the exposure of the positive charges at the surface. In the N-terminal portion of apolipoprotein B-26 of the invention,

referring to Figure 1, concentrations of lysine and arginine residues occur at positions 18-24, 87-90, 166-169, and 225-228. Cysteine residues at positions 12, 51, 61, 70, 159, 185, 218, and 234 provide  
5 opportunities for disulfide cross-links to ensure exposure of these four regions of positive charge. The receptor binding region of apoE has been determined by others to reside in the region of amino acid residues 140-160 (McClellan, J.W., et al, supra). This region  
10 contains a high concentration of arginine residues as well as two lysines.

To fit the definition, a peptide sequence need only, however, successfully compete with labelled LDL in the receptor binding assay.

15 A "positively charged domain" as used in describing a sector of an LDL receptor binding region refers to a sequence of four consecutive amino acids bearing, at neutral pH, at least three positive charges. Amino acid side chains which confer positive  
20 charges at neutral pH are arginine and lysine. Histidine residues confer approximately one-half the value of a charge at pH 7; therefore two histidine residues would be required to obtain the same charge density as that conferred by one lysine or arginine.

25 A "lipophilic" protein as used herein refers to a protein capable of binding phospholipids or other lipids such as cholesterol, cholesteryl esters, or triglycerides to form protein lipid complexes, as measured, for example, by their circular dichroic (DC)  
30 spectra (see Sparrow, J.T., et al, supra and Morrisett, et al, Ann Rev Biochem (1975) 44: 183-207). The capacity of proteins to complex with lipids has been characterized fairly extensively. It is considered by Sparrow, J.T., et al, supra, to be primarily a function

-10-

of potential to form an  $\alpha$  helical segment, a critical length of about 20 amino acids to stabilize the helix, and a minimum hydrophobicity of the non-polar face of the amphipathic helix of about -850 kcals/residue.

5 These parameters are discussed in the referenced paper.

The amphipathic  $\alpha$  helix contains a sequence of amino acids such that an  $\alpha$  helix conformation is stabilized and such that the hydrophobic side chain amino acids reside on one face of the helix while the  
10 hydrophilic and charged side chain amino acids reside on the opposite face. Thus the helix presents a hydrophilic face and hydrophobic face. This permits association of lipids with the hydrophobic face while solubilizing the entire complex through the  
15 hydrophilicity of the other.

"Operably linked" refers to a juxtaposition wherein the components are configured so as to perform their usual function. Thus, control sequences operably linked to coding sequences are capable of effecting the  
20 expression of the coding sequence.

"Control sequence" refers to a DNA sequence or sequences which are capable, when properly ligated to a desired coding sequence, of effecting its expression in hosts compatible with such sequences. Such control  
25 sequences include promoters in both procaryotic and eucaryotic hosts, and in procaryotic organisms also include ribosome binding site sequences, and, in eucaryotes, termination signals. Additional factors necessary or helpful in effecting expression may  
30 subsequently be identified. As used herein, "control sequences" simply refers to whatever DNA sequence may be required to effect expression in the particular host used.

-11-

"Cells" or "recombinant host cells" or "host cells" are often used interchangeably as will be clear from the context. These terms include the immediate subject cell, and, of course, the progeny thereof. It is understood that not all progeny are exactly identical to the parental cell, due to chance mutations or differences in environment. However, such altered progeny are included when the above terms are used.

"Peptide" and "protein" are used interchangeably herein, without regard to chain length.

"Derived from" does not necessarily refer to physical derivation. It means, in the case of peptides, that the amino acid sequences correspond to those of a referenced protein; in the case of a DNA that the base sequence corresponds to that of a referenced oligonucleotide, and so forth.

"Active substance" or "ingredient" refers to a material which exerts an intended effect on a subject organism. It may, for example, be therapeutic, regulatory, toxic, or nutritive.

#### B. Compositions of the Invention

The compositions of the invention contain

(a) a peptide sequence which is an LDL receptor binding region conjugated to an active ingredient or to a carrier for an active ingredient; or

(b) a lipophilic protein containing at least one amphipathic  $\alpha$  helical region covalently bound to a target-specific peptide; or

(c) both a peptide sequence which is an LDL receptor binding region and a lipophilic protein containing at least one amphipathic  $\alpha$  helical region covalently bound to a target specific peptide associated with an active ingredient.

-12-

### B.1 LDL Receptor Binding Region

For those embodiments which contain the peptide sequence which is an LDL receptor binding region, this sequence is covalently bound to a substance which is  
5 either itself an active ingredient in a pharmaceutical or nutrient composition or is a carrier for such active ingredients.

One embodiment of the active substance covalently bound to the LDL receptor binding region is  
10 an additional polypeptide fused to the binding region which is produced by construction, at the DNA level, of a nucleotide sequence encoding the entire fusion protein. While conjugation at the protein level is theoretically possible, it is difficult to accomplish  
15 using standard techniques since the LDL binding region itself contains the side chain amino groups and the cysteine residues ordinarily used as functional groups in binding to linkers conventionally used in formation of, for example, immunoconjugates. To the extent that  
20 peptide linkers specific for carboxy terminal groups on the LDL receptor region can be found, covalent binding at the protein level is feasible.

Any polypeptide whose amino acid sequence is known, or for which a cDNA or genomic clone exists can  
25 be used as the active ingredient. Such peptides include, for example, human growth hormone, urokinase, tumor necrosis factor,  $\alpha$ -,  $\beta$ -, and  $\gamma$ -interferons, various toxins such as ricin toxin, ricin A chain, diphtheria toxin, PAP, abrin, gelonin, and the like,  
30 insulin, various growth factors, hormones such as LHRH, and so forth. Any biologically active peptide for which it is possible to construct a DNA sequence encoding primary structure is suitable for fusion to the DNA

-13-

encoding LDL receptor binding region for the compositions of the invention.

Additional active ingredients, which are not peptides, but which are reactive or can be made reactive  
5 with the carboxyl group of the LDL receptor binding peptide may also be directly covalently linked to the binding region peptide.

Constructions in which the LDL receptor binding region is covalently bound to carrier may be effected in  
10 a similar manner to that described for attachment to a peptide active ingredient. The carrier may be the lipophilic protein bearing an amphipathic  $\alpha$  helical domain as is described in the illustration below. Additional protein carriers include human serum albumin  
15 which is especially suitable for water soluble compounds. The role of the albumins in drug transport is already well established, and the DNA sequence encoding human serum albumin (HSA) is known. Therefore, HSA encoding DNA may be fused to the LDL receptor  
20 binding region encoding DNA to produce a sequence encoding a fusion in which the HSA acts as the carrier moiety. The active ingredient in this case is generally hydrophilic, and includes such conventional drugs as Warfarin, penicillinase resistant penicillins, and  
25 sulfonamides.

The construction of a DNA sequence encoding the desired fusion protein with active peptide ingredient or carrier utilizes standard restriction endonuclease mediated cleavage techniques, techniques for  
30 modification of termini, as required, such as blunt ending, removal of nucleotide sequences, ligation of linkers, and the like, and religation of the components of the sequence. Typically, the portions of the DNA sequence are obtained either using synthetic techniques

-14-

such as those employing commercially available DNA synthesizers in combination with annealing of overlapping single strands, filling in with DNA polymerase, and ligation of portions of the sequences, or parts may be derived from cDNA or genomic clones. Methods for construction of a desired DNA sequence using such sources are by now well established in the art.

Construction of expression systems using control sequences compatible with procaryotes and eucaryotic hosts is also understood. A number of control sequences usable in expressing the coding sequence for the fusion proteins of the invention are set forth in the section on Standard Methods contained hereinbelow. This section also summarizes standard procedures for transformation and expression of these systems in a variety of exemplary hosts.

#### B.2. The Amphipathic Domain

Conversely, the amphipathic domain can be conjugated at the DNA or protein level as appropriate to other specific receptor binding moieties such as, most preferably, immunoglobulins as well as to the LDL-receptor binding region. Techniques for conjugation of proteins into conjugate immunotoxins at the protein level are well understood in the art. Commercially available linkers include the popular SPDP which results in an amide linkage to a side chain amino group at one terminus of the linker and a sulfhydryl at the other, as well as linkers forming thioethers at this terminus. Commercially available linkers having different functionalities at either terminus are now well known in the art. Among the most commonly used are SPDP, and the active esters, for example succinimidyl esters, of 6-maleimidocaproic or 2-iodoacetic acid.

-15-

The amphipathic helical region of the lipophilic protein is used as a carrier for the active ingredient in the composition. The active ingredient may simply itself be a lipid, such as, for example, 5 Intralipid, a fat soluble vitamin composition microemulsion. The microemulsion may also contain additional active ingredients, such as fat soluble drugs including vitamins and chemotherapeutic agents as is described in U.S. serial no. 765,359, filed 13 August 10 1985, assigned to the same assignee and incorporated herein by reference. Thus, the lipophilic protein serves not as an active ingredient itself, but rather as a carrier to support this active ingredient.

For those embodiments wherein the carrier 15 protein is a lipophilic protein bearing an amphipathic  $\alpha$  helical domain, the binding region may be fused either at the protein or the DNA level. Of course, if the LDL receptor binding moiety is used in this regard and fusion at the DNA level is preferred. However, for 20 conjugation of proteins having free sulfhydryl or amino terminal residues for linkage, conjugation at the protein level may also be convenient. When the binding moiety is fused to the lipophilic portion, the lipophilic portion behaves as carrier in a manner 25 described as above for a suitable active hydrophobic ingredient.

#### C. Mode of Administration

The compositions of the invention are 30 administered as pharmaceutical compositions in a conventional manner. They may be administered parenterally, for example, intravenously or subcutaneously. With proper formulation, it may be possible to administer them as oral dosages. The dosage

-16-

level is, of course, highly dependent on the nature of the active ingredient, and must be calculated based on the efficacy of this ingredient, as well as on the usual parameters relating to the nature of the subject and  
5 condition to be treated.

For parenteral administration, the delivery compositions of the invention are formulated as unit dosages in injectable form, such as an emulsion, suspension or solution, in association with at least one  
10 standard excipient, such as water, dextrose solution, saline solution, and the like. The nature of such excipients and formulations is understood by those skilled in the art.

15 D. Illustrative Embodiment

The following illustration is intended to demonstrate the method of the invention in a specific instance, but is not meant to limit it. In this illustration, the LDL receptor binding region is derived  
20 from amino acids 1-250 of apoB protein and the two lipophilic amphipathic helical domains extending through amino acid 402 of this sequence are retained. This apoB derived peptide is itself exemplary of the composition of the invention. In another illustration, this portion  
25 is fused through a tripeptide linker to the lipophilic domains of apolipoprotein AI representing the C-terminal 223 amino acids of apoAI.

30 D.1. Verification of the LDL Receptor Binding Domain

A 17 amino acid peptide representing the sequence from Thr<sub>17</sub> to Ser<sub>33</sub> of apoB as shown in Figure 1 was prepared using standard solid phase peptide synthesis with a commercial synthesizer. The product,

judged to be more than 98% pure by HPLC analysis, was used to prepare antisera.

Two mg of the peptide was dissolved in 1 mM HCl, treated with 1 mg carbodiimide for 15 min at 4°C and then mixed with 2 mg keyhole lipid hemocyanin for 12 hr at room temperature. The mixture was dialyzed against phosphate buffered saline for 24 hr at 4°C, and the mixture emulsified with Freund's complete adjuvant for injection into rabbits. Antisera obtained from the injected white New Zealand rabbits by Berkeley Antibody Company (Richmond, CA) was titered against the peptide using solid phase ELISA as described by Engvold, E., Meth Enzym (1980) 70: 365-387. High titer antisera against the 17 amino acid peptide was assayed for its ability to inhibit the binding of LDL to its receptor on intact fibroblasts according to the method of Goldstein, J.L., et al, Cell (1976) 7: 85-95 and to the receptors on bovine liver membranes (Basu, S.K., et al, PNAS (1976) 73: 3178-3182). The antisera were successful in inhibiting this binding.

In a completely similar manner, synthetic peptides representing amino acid residues 259-280, 158-186, and 217-235 of the apoB sequence as shown in Figure 1 were prepared and assayed for inhibition of LDL receptor binding activity and are found to be similarly active.

The foregoing results show that these regions of apoB, shown overlined in Figure 1, which span the positively charged sequences identified above are interactive with the LDL receptor.

#### D.2. Preparation of ApoB Encoding cDNA

An approximately  $5 \times 10^5$  member human adult liver cDNA library (where the insert size averaged 1 kb

-18-

and the inserts were ligated into the EcoRI site of  $\lambda$ gt10) was prepared by the method of Huynh, T., et al, DNA Cloning Techniques: A Practical Approach (1984), Grover, D., ed., IRL Press, Oxford. For screening, 9 x 10<sup>5</sup> plaques propagated in C600 (HFL) cells were transferred to replica nitrocellulose filters and processed as described by Seilhamer, J.J., et al, DNA (1984) 3: 309-317. The filters were prewashed for 2 hr in 3 x NaCl/Cit (1 x NaCl/Cit is 150 mM NaCl/15 mM sodium citrate, pH 7.0), 0.1% SDS at 55°C, and then prehybridized in 6 x NaCl/Cit, 200  $\mu$ g/ml denatured salmon sperm DNA, 5 x Denhardt's, 0.05% sodium pyrophosphate for 1 hr at 50°C.

A 192-fold degenerate 23 base oligonucleotide probe which encodes, taking account of codon redundancy, the first 8 amino acids of the previously determined sequence of apoB-26 was used as a probe. The probe was 5' end labelled with T<sub>4</sub> polynucleotide kinase (PL Biochemicals) and  $\gamma$  32-PATG, added to the filters and incubated for 14 hr at 50°C. The filters were washed twice at room temperature in 5 x NaCl/Cit, 0.1% SDS, 0.05% sodium pyrophosphate for 15 min and once at 50°C for 20 min, dried and autoradiographed with intensifying screens.

One positive plaque, designated LB25-1, was purified and the cDNA insert was subcloned in both orientations into M13/mp8 for sequencing. The nucleotide sequence of this 970 bp insert is shown in Figure 2, along with the deduced amino acid sequence of that reading frame which agrees with the amino terminal sequence determined from the B-26 peptide. The sequence of LB25-1 contains an open reading frame extending 800 nucleotides downstream encoding 30 kd of protein, and analysis of the predicted protein sequence directly

-19-

upstream suggests the presence of a hydrophobic signal sequence preceded by a methionine residue. The EcoRI insert was subcloned into pBR322 to obtain pB25-1 for amplification. pB25-1 thus contains some 5'

5 untranslated region, the signal sequence, and the first 266 amino acids of the mature protein.

Additional portions of the apoB encoding sequence were obtained using linearized denatured pB25-1 insert as probe. One such segment, designated LB2A-2  
10 was purified and partially sequenced as described above. The approximately 2 kb insert was then subcloned into pBR322 to obtain pB25-2 to provide additional downstream apoB sequence.

15 D.3. Construction of DNA Encoding ApoB LDL Binding Region/Lipophilic Region

DNA encoding the LDL binding region, which is contained in the 266 amino acid encoding sequence of the pB25-1 insert and a lipophilic region containing two  
20 domains of amphipathic  $\alpha$  helical sequences which is contained in the sequence extending from 267-433 of the mature protein was prepared from pB25-1 and pB2A-2.

The EcoRI insert from pB25-1 was removed and treated with BclI (see Figure 2); the resulting 806 bp  
25 EcoRI/BclI fragment was isolated. pB2A-2, containing the 2 kb insert with additional apoB sequence, was treated with EcoRI, the insert isolated, and the insert then cut with BclI. The resulting 1745 bp BclI/EcoRI fragment was isolated. The two isolated fragments were  
30 ligated into an EcoRI treated, BAPped, pUC8 cloning vector.

The resulting 2550 bp insert fragment was again isolated by digestion with EcoRI followed by purification using agarose gel electrophoresis. The

-20-

purified insert was treated with XhoII and the EcoRI/XhoII fragment containing the 5' region of the apoB cDNA subcloned into EcoRI/BamHI digested pUC8.

Digestion of the resultant with EcoRI and SalI removes the desired DNA encoding 402 amino acids of mature protein including the signal sequence and 5' untranslated region for positioning in a suitable expression system or for the apoAI fusion described below. This DNA encodes a protein containing both LDL receptor binding and lipophilic functions and is shown, along with the deduced sequence in Figure 3. Accordingly, this DNA may be ligated into a suitable expression system to provide one embodiment of the drug delivery compositions of the invention.

15

#### D.4. Construction of DNA Encoding ApoB/ApoAI Fusion Protein

To provide additional amphipathic helical regions, the C-terminal 220+ amino acids of apoAI was provided using genomic DNA.

First, the EcoRI/SalI apoB encoding insert of D.3 was ligated to a synthetic oligonucleotide designed as a SalI/SmaI linker region. The nucleotide sequence of the linker and the relevant amino acid sequence encoded are shown in Figure 4a. The linker was constructed from oligomers 1-6 shown in the figure using standard annealing and ligation techniques. The linker DNA provides codons for an Arg-Arg-Gln tripeptide linker and tracks the DNA sequence extending through the end of the apoAI third exon, into the intron to the SmaI site, i.e., the last 3 2/3 codons in exon 3 of the apoAI gene, the exon/intron boundary, and 44 nucleotides of intron 3. The ligated fragment was digested with EcoRI and SmaI, and subcloned into EcoRI/SmaI digested pUC8.

To add the apoAI sequences, the resulting plasmid was partially digested with PstI (which cuts just downstream of SmaI) and completely with SmaI and ligated to the StuI/HgaI (apoAI fragment 1) and HgaI/PstI (apoAI fragment 2) fragments, described below, to obtain a plasmid containing the DNA sequence shown in Figure 5. Figure 5 shows the DNA sequence encoding the apoB/AI fusion, along with the deduced amino acid sequence. This plasmid, designated pB/AI was used as a source for the DNA sequence encoding the fusion protein containing LDL receptor binding region and some lipophilic region from apoB, and additional lipophilic region from AI.

The apoAI gene fragments were prepared as follows and are shown in Figure 4b. pPSAI was isolated from the  $\lambda$  genomic clone LA1.12 as disclosed by Seilhamer, et al, (1984) supra, and contains most of the apoAI gene. The plasmid was cut with PstI and the 2.2 kb fragment isolated. The purified insert was then cut with HgaI and the HgaI/PstI fragment (fragment 2) isolated. Another sample of the purified insert was also cut with StuI and HgaI and StuI/HgaI fragment (fragment 1) isolated.

#### 25 D.5. Construction of Expression Vectors and Expression of B5' and B/AI Fusions

The DNA segment encoding the apoB LDL-receptor binding and lipophilic regions (B5') is used to produce the recombinant protein in a variety of hosts. However, expression in mammalian systems is favored as the host is capable of post translational processing analogous to that experienced by the natively produced protein. Either uninterrupted or genomic sequences including the

-22-

foregoing DNA encoding the B/AI fusion may be used, as the host is also capable of processing introns.

In an illustrative embodiment, DNAs encoding the compositions of the invention as prepared in ¶D.3 or D.4 are excised with EcoRI and SalI, or with EcoRI and PstI respectively, provided with EcoRI linkers at the SalI or PstI sites and ligated into pHS1, a host expression vector described below. The resulting expression vectors are transformed into Chinese Hamster Ovary cells and the cells induced as described below. High producing clones are identified by screening with appropriate antibodies. For the B5' fragment of ¶D.3, antibodies raised against this fragment or the LDL-receptor binding regions thereof are used. For the apoB/AI fusion clones showing positive screening against both anti-apoB and anti-apoAI are selected.

D.5.a. Construction of pHS1, the Host Expression Vector

The plasmid pHS1 contains 840 bp of the hMT-II sequence from p84H (Karin, M., et al, Nature (1982) 299: 297-802) which spans from the HindIII site at position -765 of the hMT-II gene to the BamHI cleavage site at base + 70. Plasmid p84H was digested to completion with BamHI, treated with exonuclease Bal-31 to remove terminal nucleotides, and then digested with HindIII. The desired 840 bp fragment was ligated into pUC8 (Vieira, J., et al, Gene (1982) 19: 259-268) which had been opened with HindIII and HincII digestion. The ligation mixture was transformed into E. coli HB101 to Amp<sup>R</sup>, and one candidate plasmid, designated pHS1, was isolated and sequenced by dideoxy sequencing. pHS1 contains the hMT-II control sequences upstream of a polylinker containing convenient restriction sites

including an EcoRI site to place an inserted DNA into operable linkage with the promoter.

D.5.b. Construction of Expression Vectors

5       The EcoRI/EcoRI B5' or B/AI fusion encoding fragment, prepared as above, is ligated into EcoRI digested pHS1 and the ligation mixture transformed into E. coli MC1061 to Amp<sup>R</sup>. Successful transformants are screened by restriction analysis, and a strain  
10       containing the desired plasmid, pMT-B5' or pMT-B/AI, is further propagated to prepare quantities of plasmid DNA.

Production of B5' and the B/AI Fusion by Mammalian Recombinants

15       Chinese hamster ovary (CHO)-K1 cells are grown on medium composed of a 1:1 mixture of F12 medium and DME medium with 12% fetal calf serum. The competent cells are co-transformed with pMT-B5' or pMT-B/AI and pSV2:NEO (Southern, P., et al, J Mol Appl Genet (1982)  
20       1: 327-341). pSV2:NEO contains a functional gene conferring resistance to the neomycin analog G418. In the transformation, 500 ng of pSV2-NEO and 5 µg of pMT-B5' or pMT-B/AI are applied to a 16 mm dish of cells in a calcium phosphate-DNA co-precipitate according to  
25       the protocol of Wigler, M., et al, Cell (1979) 16: 777-785, with the inclusion of a two minute "shock" with 15% glycerol after four hours of exposure to the DNA. A day later, the cells are subjected to 1 mg/ml G418 to provide a pool of G418-resistant colonies.

30       Successful transformants, also having a stable inheritance of pMT-B5' or pMT-B/AI, are plated at low density for purification of clonal isolates. Small amounts of these isolates are grown in multi-well plates after exposure to  $10^{-4}$  M zinc chloride for convenient

assay of B5' or B/AI fusion protein production. These determinations are made by standard radio-immunoassays or ELISA against the antisera prepared against apoB and in the case of pMT-B/AI also against apoAI. Clonal  
5 isolates which produce high titers with respect to the appropriate antisera are selected.

The cells, which have been shown to produce the B5' or B/AI fusion under suitable conditions, are seeded at 1/10 confluency in basal medium supplemented with 10%  
10 fetal calf serum, incubated overnight, and then induced for protein production by addition of zinc chloride in the concentration range of  $1 \times 10^{-4}$  M to  $3 \times 10^{-4}$  M. The levels of the B5' or B/AI fusion rise for 7-10 days, under optimal inducing conditions,  $2 \times 10^{-4}$  M  
15  $\text{ZnCl}_2$ .

If desired, the B5' or B/AI fusion secreted into the medium can be further purified according to the procedures set forth for the amphipathic helix-containing protein summarized below by standard methods  
20 known in the art.

Proteins containing an amphipathic helix can be readily complexed with lipids. The resulting lipoprotein complex can be separated readily from nonlipid binding proteins on the basis of either size or  
25 density as follows.

The lipoprotein complex is a particulate complex composed of lipid particles -- either emulsion, liposomal, or lipid-disc particles -- with associated apolipoprotein. The lipoprotein particles are in the  
30 general size range of between about 0.01 to 0.55 $\mu$ , and as such, can be readily separated from smaller, soluble proteins by particle exclusion in molecular sieve chromatography. The molecular-sieve material is preferably selected to exclude material in the range

500,000 daltons or greater, insuring that essentially all of noncomplexed proteins will be included (retarded) by passage through the chromatography column bed. The larger lipoprotein complex, which is excluded, will  
5 elute in the void volume of the column. The eluted fractions can be monitored, conventionally, by UV spectroscopy or the like. Preferred molecular-sieve material includes Agarose 10%, whose exclusion size is about 1,000,000 daltons.

10 For large-scale separation, the lipoprotein complex is preferably separated from nonlipid binding peptides by flotation. Typically this is done by adjusting the density of the peptide-mixture medium to between about 1.06-1.21 by addition of a mono- or  
15 disaccharide, such as sucrose, or a salt, such as KCl. The medium is then placed in a centrifuge tube, overlaid with a lower density medium, and centrifuged until the complex has concentrated at the top of the tube. The lipoprotein material removed from the  
20 gradient may be resuspended in the same density medium and separated a second time, to further remove nonlipid-binding proteins.

The lipoprotein complex will separate on its own by flotation if allowed to stand undisturbed in a  
25 medium of sufficient density. This flotation procedure is, of course, advantageous in large-scale preparation, since both chromatography and centrifugation steps are eliminated.

The apolipoprotein associated with the  
30 separated lipoprotein complex may be further purified, as necessary, by delipidating the complex and purifying the peptide by conventional purification methods.

Complexing may be done as follows: INTRALIPID, obtained from Cutter Labs (Berkeley, CA) can be used as

-26-

the substrate to bind the apolipoprotein ligand. INTRALIPID is an artificial lipid emulsion composed of soybean triacylglycerol and egg lecithins. The mesophase, or phospholipid-rich portion of the emulsion is removed by ultracentrifugal flotation in a discontinuous sucrose gradient to obtain the triglyceride-rich emulsion which is at the top of the gradient. The apolipoprotein is incubated with purified INTRALIPID and centrifuged as described for INTRALIPID to obtain the top fraction. The apolipoprotein is recovered by delipidation if desired.

#### E. Standard Methods

Most of the techniques which are used to transform cells, construct vectors, extract messenger RNA, prepare cDNA libraries, and the like are widely practiced in the art, and most practitioners are familiar with the standard resource materials which describe specific conditions and procedures. However, for convenience, the following paragraphs may serve as a guideline.

##### E.1. Hosts and Control Sequences

Both procaryotic and eucaryotic systems may be used to express the DNA sequences encoding the recombinant proteins useful in the invention; procaryotic hosts are, of course, the most convenient for cloning procedures. Procaryotes most frequently are represented by various strains of E. coli; however, other microbial strains may also be used. Plasmid vectors which contain replication sites and control sequences derived from a species compatible with the host are used; for example, E. coli is typically transformed using derivatives of pBR322, a plasmid

derived from an E. coli species by Bolivar, et al, Gene (1977) 2:95. pBR322 contains genes for ampicillin and tetracycline resistance, and thus provides additional markers which can be either retained or destroyed in constructing the desired vector. Commonly used procaryotic control sequences which are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding site sequences, include such commonly used promoters as the beta-lactamase (penicillinase) and lactose (lac) promoter systems (Chang, et al, Nature (1977) 198:1056 and the tryptophan (trp) promoter system (Goeddel, et al Nucleic Acids Res (1980) 8:4057 and the lambda derived P<sub>L</sub> promoter and N-gene ribosome binding site (Shimatake, et al, Nature (1981) 292:128).

In addition to bacteria, eucaryotic microbes, such as yeast, may also be used as hosts. Laboratory strains of Saccharomyces cerevisiae, Baker's yeast, are most used although a number of other strains are commonly available. Vectors employing, for example, the 2  $\mu$  origin of replication of Broach, J. R., Meth Enz (1983) 101:307, or other yeast compatible origins of replications (see, for example, Stinchcomb, et al, Nature (1979) 282:39, Tschempe, et al, Gene (1980) 10:157 and Clarke, L. et al, Meth Enz (1983) 101:300) may be used. Control sequences for yeast vectors include promoters for the synthesis of glycolytic enzymes (Hess, et al, J Adv Enzyme Reg (1968) 7:149; Holland, et al, Biochemistry (1978) 17:4900). Additional promoters known in the art include the promoter for 3-phosphoglycerate kinase (Hitzeman, et al, J Biol Chem (1980) 255:2073), and those for other glycolytic enzymes. Other promoters, which have the additional advantage of transcription controlled by

growth conditions are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and enzymes responsible for maltose and galactose utilization. It is also believed terminator sequences are desirable at the 3' end of the coding sequences. Such terminators are found in the 3' untranslated region following the coding sequences in yeast-derived genes.

It is also, of course, possible to express genes encoding polypeptides in eucaryotic host cell cultures derived from multicellular organisms. See, for example, Axel, et al, U.S. patent 4,399,216. These systems have the additional advantage of the ability to splice out introns and thus can be used directly to express genomic fragments. Useful host cell lines include VERO and HeLa cells, and Chinese hamster ovary (CHO) cells. Expression vectors for such cells ordinarily include promoters and control sequences compatible with mammalian cells such as, for example, the commonly used early and late promoters from Simian Virus 40 (SV 40) (Fiers, et al, Nature (1978) 273:113), or other viral promoters such as those derived from polyoma, Adenovirus 2, bovine papiloma virus, or avian sarcoma viruses. The controllable promoter, hMT-II (Karin, M., et al, Nature (1982) 299:797-802) may also be used. General aspects of mammalian cell host system transformations have been described by Axel (supra). It now appears, also that "enhancer" regions are important in optimizing expression; these are, generally, sequences found upstream or downstream of the promoter region in non-coding DNA regions. Origins of replication may be obtained, if needed, from viral sources. However, integration into the chromosome is a common mechanism for DNA replication in eucaryotes.

### E.2. Transformations

Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described by Cohen, S.N., Proc Natl Acad Sci (USA) (1972) 69:2110, or the  $\text{RbCl}_2$  method described in Maniatis, T., et al, Molecular Cloning: A Laboratory Manual (1982) Cold Spring Harbor Press, p. 254 may be used for procaryotes or other cells which contain substantial cell wall barriers. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology (1978) 52:546, optionally as modified by Wigler, M., et al, Cell (1979) 16:777-785 may be used. Transformations into yeast may be carried out according to the method of Van Solingen, P., et al, J Bact (1977) 130:946 or of Hsiao, C.L., et al, Proc Natl Acad Sci (USA) (1979) 76:3829.

### 20 E.3. Vector Construction

Construction of suitable vectors containing the desired coding and control sequences employs standard ligation and restriction techniques which are well understood in the art. Isolated plasmids, DNA sequences, or synthesized oligonucleotides are cleaved, tailored, and religated in the form desired.

Site specific DNA cleavage is performed by treating with the suitable restriction enzyme (or enzymes) under conditions which are generally understood in the art, and the particulars of which are specified by the manufacturer of these commercially available restriction enzymes. See, e.g., New England Biolabs, Product Catalog. In general, about 1  $\mu\text{g}$  of plasmid or DNA sequence is cleaved by one unit of enzyme in about

-30-

20 µl of buffer solution; in the examples herein, typically, an excess of restriction enzyme is used to insure complete digestion of the DNA substrate. Incubation times of about one hour to two hours at about 37°C are workable, although variations can be tolerated. After each incubation, protein is removed by extraction with phenol/chloroform, and may be followed by ether extraction, and the nucleic acid recovered from aqueous fractions by precipitation with ethanol. If desired, size separation of the cleaved fragments may be performed by polyacrylamide gel or agarose gel electrophoresis using standard techniques. A general description of size separations is found in Methods in Enzymology (1980) 65:499-560.

Restriction cleaved fragments may be blunt ended by treating with the large fragment of E. coli DNA polymerase I (Klenow) in the presence of the four deoxynucleotide triphosphates (dNTPs) using incubation times of about 15 to 25 min at 20 to 25°C in 50 mM Tris pH 7.6, 50 mM NaCl, 6 mM MgCl<sub>2</sub>, 6 mM DTT and 5-10 µM dNTPs. The Klenow fragment fills in at 5' sticky ends but chews back protruding 3' single strands, even though the four dNTPs are present. If desired, selective repair can be performed by supplying only one of the, or selected, dNTPs within the limitations dictated by the nature of the sticky ends. After treatment with Klenow, the mixture is extracted with phenol/chloroform and ethanol precipitated. Treatment under appropriate conditions with S1 nuclease or Bal-31 results in hydrolysis of any single-stranded portion.

Synthetic oligonucleotides are prepared by the method of Efimov, V.A., et al (Nucleic Acids Res (1982) 6875-6894), and can be prepared using commercially available automated oligonucleotide synthesizers.

-31-

Kinasing of single strands prior to annealing or for labeling is achieved using an excess, e.g., approximately 10 units of polynucleotide kinase to 1 nmole substrate in the presence of 50 mM Tris, pH 7.6,

- 5 10 mM  $MgCl_2$ , 5 mM dithiothreitol, 1-2 mM ATP, 1.7 pmoles  $\gamma^{32}P$ -ATP (2.9 mCi/mmole), 0.1 mM spermidine, 0.1 mM EDTA.

Ligations are performed in 15-50  $\mu$ l volumes under the following standard conditions and

- 10 temperatures: 20 mM Tris-Cl pH 7.5, 10 mM  $MgCl_2$ , 10 mM DTT, 33  $\mu$ g/ml BSA, 10 mM-50 mM NaCl, and either 40  $\mu$ M ATP, 0.01-0.02 (Weiss) units T4 DNA ligase at 0°C (for "sticky end" ligation) or 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at 14°C (for "blunt end" ligation).  
15 Intermolecular "sticky end" ligations are usually performed at 33-100  $\mu$ g/ml total DNA concentrations (5-100 nM total end concentration). Intermolecular blunt end ligations (usually employing a 10-30 fold molar excess of linkers) are performed at 1  $\mu$ M total  
20 ends concentration.

In vector construction employing "vector fragments", the vector fragment is commonly treated with bacterial alkaline phosphatase (BAP) or calf intestinal alkaline phosphatase (CIP) in order to remove the 5'

- 25 phosphate and prevent religation of the vector. Digestions are conducted at pH 8 in approximately 150 mM Tris, in the presence of  $Na^+$  and  $Mg^{+2}$  using about 1 unit of BAP or CIP per  $\mu$ g of vector at 60° for about one hour. In order to recover the nucleic acid  
30 fragments, the preparation is extracted with phenol/chloroform and ethanol precipitated. Alternatively, religation can be prevented in vectors which have been double digested by additional restriction enzyme digestion of the unwanted fragments.

-32-

For portions of vectors derived from cDNA or genomic DNA which require sequence modifications, site specific primer directed mutagenesis is used. This is conducted using a primer synthetic oligonucleotide complementary to a single stranded phage DNA to be mutagenized except for limited mismatching, representing the desired mutation. Briefly, the synthetic oligonucleotide is used as a primer to direct synthesis of a strand complementary to the phage, and the resulting double-stranded DNA is transformed into a phage-supporting host bacterium. Cultures of the transformed bacteria are plated in top agar, permitting plaque formation from single cells which harbor the phage.

Theoretically, 50% of the new plaques will contain the phage having, as a single strand, the mutated form; 50% will have the original sequence. The resulting plaques are hybridized with kinased synthetic primer at a temperature which permits hybridization of an exact match, but at which the mismatches with the original strand are sufficient to prevent hybridization. Plaques which hybridize with the probe are then picked, cultured, and the DNA recovered.

#### 25 E.4. Verification of Construction

In the constructions set forth below, correct ligations for plasmid construction are confirmed by first transforming E. coli strain MC1061 obtained from Dr. M. Casabadian (Casabadian, M., et al, J Mol Biol (1980) 138:179-207) or other suitable host with the ligation mixture. Successful transformants are selected by ampicillin, tetracycline or other antibiotic resistance or using other markers depending on the mode of plasmid construction, as is understood in the art.

Plasmids from the transformants are then prepared according to the method of Clewell, D.B., et al, Proc Natl Acad Sci (USA) (1969) 62:1159, optionally following chloramphenicol amplification (Clewell, D.B., J Bacteriol (1972) 110:667). The isolated DNA is analyzed by restriction and/or sequenced by the dideoxy method of Sanger, F., et al, Proc Natl Acad Sci (USA) (1977) 74:5463 as further described by Messing, et al, Nucleic Acids Res (1981) 9:309, or by the method of Maxam, et al, Methods in Enzymology (1980) 65:499.

#### E.5. cDNA or Genomic Library Production

Human genomic libraries are constructed in  $\lambda$  phage as is known in the art. See, e.g., Maniatis, T., et al, Cell (1978) 15: 687-701. cDNA libraries can be prepared as described above in  $\lambda$ gt11 phage, or double-stranded cDNA synthesized from mRNA isolated using standard techniques can be prepared for insertion into a plasmid vector such as pBR322 using homopolymeric tailing mediated by calf thymus terminal transferase (Sutcliffe, J.G., Nucleic Acid Res (1978) 5:2721-2732). First strand cDNA is synthesized by the RNA-dependent DNA polymerase from Avian Myeloblastosis Virus, by priming with oligo (dT) 12-18 on 5  $\mu$ g mRNA. The RNA template is then liberated from the nascent DNA strand by denaturation at 100°C for 5 min, followed by chilling on ice. Second strand DNA is synthesized by using the large fragment of DNA polymerase I of E. coli, relying on self-priming at the 3'-end of the first strand molecule, thereby forming a double-stranded hairpin DNA. These molecules are blunt-ended at the open-ended termini, and the hairpin loop is cleaved open with S1 nuclease from Aspergillus oryzae. S1 nuclease digestion of the double-stranded cDNA takes place in 300 mM NaCl,

-34-

30 mM NaOAc, pH 4.5, 3 mM  $\text{ZnCl}_2$  for 30 min at 37°C with 600 units enzyme. The cDNA is extracted with phenol:chloroform, and small oligonucleotides are removed by three ethanol precipitations in the presence of ammonium acetate. This is done as follows: a half volume of 7.5 M ammonium acetate and two volumes ethanol are added to the cDNA solution, which is precipitated at -70°C. The blunt-ended, double-stranded cDNA is then fractionated by size using gel filtration through a column (0.3 x 14 cm) Sepharose 4B (Pharmacia Fine Chemicals, Piscataway, NJ) or by ultracentrifugation in 5-20% glycerol gradient followed by fractionation of the gradient. cDNA roughly greater than the desired length, e.g., 300 base pairs is retained and recovered by precipitation with 70% ethanol. Short (10-30 nucleotides) polymeric tails of deoxycytosine are added to the 3' termini of the cDNA using a reaction containing 0.2 M potassium cacodylate, 25 mM Tris, pH 6.9, 2 mM dithiothreitol, 0.5 mM  $\text{CoCl}_2$ , 200 mM cDTP, 400 µg/ml BSA, and 40 units calf thymus terminal deoxynucleotide transferase for 5 min at 22°C. The reaction is extracted with phenol:chloroform, and small oligonucleotides are removed with three ethanol precipitations in the presence of ammonium acetate.

The tailed cDNA is annealed with a host vector such as pBR322 which has been cleaved with, for example, PstI and tailed with oligo dG. In one operable embodiment 2.5 µg pBR322-dG DNA is annealed with the cDNA at a vector concentration of 5 µg/ml, and the hybrids are transferred into E. coli MC1061 by the  $\text{CaCl}_2$ -treatment described by Casabadan, M., et al, Mol Biol (1980) 138:179-207.

#### E.6. Probing cDNA or Genomic Libraries

cDNA or genomic libraries may be screened if desired using the colony or plaque hybridization procedures. Colonies or plaques are replicated onto  
5 duplicate nitrocellulose filter papers (S & S type BA-85) and colonies are allowed to grow at 37°C for 14-16 hr on L agar containing 15 µg/ml tetracycline. The colonies are lysed with 10% SDS and the DNA is fixed to the filter by sequential treatment for 5 min with 500  
10 mM NaOH/1.5 M NaCl, then 0.5 M Tris HCl(pH 8.0)/1.5 M NaCl followed by 2 x standard saline citrate (SSC). Filters are air dried and baked at 80°C for 2 hr.

For nick-translated probe, the duplicate filters are prehybridized at 42°C for 16-18 hr with 10  
15 ml per filter of DNA hybridization buffer (50% formamide (40% formamide if reduced stringency), 5 x SSC, pH 7.0, 5x Denhardt's solution (polyvinylpyrrolidone, plus Ficoll and bovine serum albumin; 1 x = 0.02% of each), 50 mM sodium phosphate buffer at pH 7.0, 0.2% SDS, 50  
20 µg/ml yeast tRNA, and 50 µg/ml denatured and sheared salmon sperm DNA).

Samples are hybridized with nick-translated DNA probes at 42°C for 12-36 hr for homologous species and 37°C for heterologous species contained in 5 ml of this  
25 same DNA hybridization buffer. The filters are washed two times for 30 min, each time at 50°C, in 0.2 x SSC, 0.1% SDS for homologous species hybridization, and at 50°C in 3 x SSC, 0.1% SDS for heterologous species hybridization. Filters are air dried and  
30 autoradiographed for 1-3 days at -70°C.

For synthetic (15-30 mer) oligonucleotide probes, the duplicate filters are prehybridized at 42°C for 2-8 hr with 10 ml per filter of oligo-hybridization buffer (6 x SSC, 0.1% SDS, 1 mM EDTA, 5x Denhardt's,

-36-

0.05% sodium pyrophosphate and 50 µg/ml denatured and sheared salmon sperm DNA).

The samples are hybridized with kinased oligonucleotide probes of 15-30 nucleotides under conditions which depend on the composition of the oligonucleotide. Typical conditions employ a temperature of 30-42°C for 24-36 hr with 5 ml/filter of this same oligo-hybridization buffer containing probe. The filters are washed two times for 15 min at 23°C, each time with 6 x SSC, 0.1% SDS and 50 mM sodium phosphate buffer at pH 7, then are washed once for 2 min at the calculated hybridization temperature with 6 x SSC and 0.1% SDS, air dried, and are autoradiographed at -70°C for 2 to 3 days.

If the amino acid sequence of the desired protein or nucleotide sequence encoding it in mRNA is known, the DNA for insertion into the host vectors of the invention may be obtained either by synthetic means, or, if vectors containing such sequences are on deposit or available, by cloning such vectors. For synthesis of the coding sequences, alternating sense and anti-sense overlapping single stranded oligonucleotides are prepared, and the alternating sense and anti-sense single stranded portions filled in enzymatically by treating with DNA polymerase and dNTPs. The oligomers are prepared by the method of Efimov, V.A., et al (Nucleic Acids Res (1982) 6875-6894), and can be prepared using commercially available automated oligonucleotide synthesizers. Kinasing of single strands prior to annealing or for labeling is achieved using an excess, e.g., approximately 10 units of polynucleotide kinase to 1 nmole substrate in the presence of 50 mM Tris, pH 7.6, 10 mM MgCl<sub>2</sub>, 5 mM

-37-

dithiothreitol, 1-2 mM ATP, 1.7 pmoles  $\gamma$ 32P-ATP (2.9 mCi/mmole), 0.1 mM spermidine, 0.1 mM EDTA.

E.7. Hosts Exemplified

5           Host strains used in cloning and expression herein are as follows:

          For cloning and sequencing, and for expression of construction under control of most bacterial promoters, E. coli strain MC1061 was used.

10           The cells used for mammalian expression are Chinese hamster ovary (CHO) cells.

15

20

25

30

-38-

Claims

1. A composition for delivery of an active  
5 ingredient to a vertebrate subject which comprises a  
target specific peptide comprising an LDL receptor-  
binding region covalently bound to a substance selected  
from an active ingredient and a carrier for an active  
ingredient.
- 10 2. The composition of claim 1 wherein the  
LDL-receptor binding region is derived from  
apolipoproteins B or E or wherein the LDL-receptor  
binding region consists essentially of amino acids 1-300  
15 of apolipoprotein B-26 (apoB-26) or a subfragment  
thereof wherein the subfragment contains an amino acid  
sequence selected from the group consisting of amino  
acids 18-24, 87-90, 166-169, and 225-228 of the apoB-26  
sequence shown in Figure 1.
- 20 3. The composition of claim 1 wherein the  
carrier is a lipophilic protein containing at least one  
amphipathic helix domain.
- 25 4. The composition of claim 3 wherein the  
lipophilic protein is derived from apolipoprotein B, E,  
CII, CIII, AII, or AIV, or from apolipoprotein AI  
(apoAI), or comprises two amphipathic helix domains of  
apoB-26.
- 30 5. A composition for delivery of an active  
ingredient to a vertebrate subject which comprises a  
target specific peptide covalently bound to a lipophilic  
protein containing at least one amphipathic helix domain.

-39-

6. The composition of claim 5 wherein the target specific peptide is an immunoglobulin or comprises an LDLH receptor binding region.

5           7. The composition of claim 5 which further includes a lipid emulsion.

8. A recombinant DNA sequence which encodes a fusion protein, wherein the fusion protein comprises a  
10 target specific peptide sequence and a lipophilic peptide sequence which contains at least one amphipathic helix domain.

9. The DNA sequence of claim 8 wherein the  
15 target specific peptide sequence is an LDL-receptor binding sequence or an immunoglobulin.

10. An expression system for a fusion protein comprising a target specific peptide sequence and a  
20 lipophilic peptide sequence, which expression system comprises the DNA sequence of claim 8 operably linked to control sequences.

11. A recombinant host cell transformed with  
25 the DNA sequence of claim 8 or the expression system of claim 10.

12. A method to produce a fusion protein comprising a target specific peptide sequence and a  
30 lipophilic peptide sequence, which method comprises culturing the cells of claim 11.

-40-

13. A fusion protein comprising a target specific peptide sequence and a lipophilic peptide sequence produced by the method of claim 12.

5

10

15

20

25

30

FIG. 1

— **1998** — **1999** — **2000** — **2001** — **2002** — **2003** — **2004** — **2005** — **2006** — **2007** — **2008** — **2009** — **2010** — **2011** — **2012** — **2013** — **2014** — **2015** — **2016** — **2017** — **2018** — **2019** — **2020** — **2021** — **2022** — **2023** — **2024** — **2025** — **2026** — **2027** — **2028** — **2029** — **2030** — **2031** — **2032** — **2033** — **2034** — **2035** — **2036** — **2037** — **2038** — **2039** — **2040** — **2041** — **2042** — **2043** — **2044** — **2045** — **2046** — **2047** — **2048** — **2049** — **2050** — **2051** — **2052** — **2053** — **2054** — **2055** — **2056** — **2057** — **2058** — **2059** — **2060** — **2061** — **2062** — **2063** — **2064** — **2065** — **2066** — **2067** — **2068** — **2069** — **2070** — **2071** — **2072** — **2073** — **2074** — **2075** — **2076** — **2077** — **2078** — **2079** — **2080** — **2081** — **2082** — **2083** — **2084** — **2085** — **2086** — **2087** — **2088** — **2089** — **2090** — **2091** — **2092** — **2093** — **2094** — **2095** — **2096** — **2097** — **2098** — **2099** — **2100** — **2101** — **2102** — **2103** — **2104** — **2105** — **2106** — **2107** — **2108** — **2109** — **2110** — **2111** — **2112** — **2113** — **2114** — **2115** — **2116** — **2117** — **2118** — **2119** — **2120** — **2121** — **2122** — **2123** — **2124** — **2125** — **2126** — **2127** — **2128** — **2129** — **2130** — **2131** — **2132** — **2133** — **2134** — **2135** — **2136** — **2137** — **2138** — **2139** — **2140** — **2141** — **2142** — **2143** — **2144** — **2145** — **2146** — **2147** — **2148** — **2149** — **2150** — **2151** — **2152** — **2153** — **2154** — **2155** — **2156** — **2157** — **2158** — **2159** — **2160** — **2161** — **2162** — **2163** — **2164** — **2165** — **2166** — **2167** — **2168** — **2169** — **2170** — **2171** — **2172** — **2173** — **2174** — **2175** — **2176** — **2177** — **2178** — **2179** — **2180** — **2181** — **2182** — **2183** — **2184** — **2185** — **2186** — **2187** — **2188** — **2189** — **2190** — **2191** — **2192** — **2193** — **2194** — **2195** — **2196** — **2197** — **2198** — **2199** — **2200** — **2201** — **2202** — **2203** — **2204** — **2205** — **2206** — **2207** — **2208** — **2209** — **2210** — **2211** — **2212** — **2213** — **2214** — **2215** — **2216** — **2217** — **2218** — **2219** — **2220** — **2221** — **2222** — **2223** — **2224** — **2225** — **2226** — **2227** — **2228** — **2229** — **2230** — **2231** — **2232** — **2233** — **2234** — **2235** — **2236** — **2237** — **2238** — **2239** — **2240** — **2241** — **2242** — **2243** — **2244** — **2245** — **2246** — **2247** — **2248** — **2249** — **2250** — **2251** — **2252** — **2253** — **2254** — **2255** — **2256** — **2257** — **2258** — **2259** — **2260** — **2261** — **2262** — **2263** — **2264** — **2265** — **2266** — **2267** — **2268** — **2269** — **2270** — **2271** — **2272** — **2273** — **2274** — **2275** — **2276** — **2277** — **2278** — **2279** — **2280** — **2281** — **2282** — **2283** — **2284** — **2285** — **2286** — **2287** — **2288** — **2289** — **2290** — **2291** — **2292** — **2293** — **2294** — **2295** — **2296** — **2297** — **2298** — **2299** — **2300** — **2301** — **2302** — **2303** — **2304** — **2305** — **2306** — **2307** — **2308** — **2309** — **2310** — **2311** — **2312** — **2313** — **2314** — **2315** — **2316** — **2317** — **2318** — **2319** — **2320** — **2321** — **2322** — **2323** — **2324** — **2325** — **2326** — **2327** — **2328** — **2329** — **2330** — **2331** — **2332** — **2333** — **2334** — **2335** — **2336** — **2337** — **2338** — **2339** — **2340** — **2341** — **2342** — **2343** — **2344** — **2345** — **2346** — **2347** — **2348** — **2349** — **2350** — **2351** — **2352** — **2353** — **2354** — **2355** — **2356** — **2357** — **2358** — **2359** — **2360** — **2361** — **2362** — **2363** — **2364** — **2365** — **2366** — **2367** — **2368** — **2369** — <

**FIG. 1 (con't)**

CTG CCG CTG AGG AG CCG CCG CAG CCAG CCAGGG CCG CAG AGG CCG AGG CCAGG CCG CAG CCG CA 62

GGAG CCG CCG CCG CAG CAG CTGG CGATGG ACCCG CCG AGG CCG CCG CTG CTGG CCG CTG CTGG CCG CTG CCT 131  
 MET Asp Pro Pro Arg Pro Ala Leu Leu Ala Leu Leu Ala Leu Pro  
 (-20)

CGC CTG CTG CTG CTG CTG CTGG CCG GCG CCAGGG CCG AAG AGG AAATG CTGG AAAATG TCAG CCTGG TC 200  
 Ala Leu Leu Leu Leu Leu Ala Gly Ala Arg Ala Glu Glu Glu MET Leu Glu Asn Val Ser Leu Val  
 (-10) (1) (10)

TGT CCA AAAGATG CG ACCCG ATT CAAG CAC CTCCGGAAG TAC ACATACAACTATGAGG CTGAGAGT TCC 269  
 Cys Pro Lys Asp Ala Thr Arg Phe Lys His Leu Arg Lys Tyr Thr Tyr Asn Tyr Glu Ala Glu Ser Ser  
 (20) (30)

AGTGGAGTCCCTGGGACTGCTGATTCAAGAAGTGCCACCAGGATCAACTGCAAGGTTGAGCTGGAGGTT 338  
 Ser Gly Val Pro Gly Thr Ala Asp Ser Arg Ser Ala Thr Arg Ile Asn Cys Lys Val Glu Leu Glu Val  
 (40) (50)

CCCCAGCTCTGCAGCTTCATCCTGAAGACCAGCCAGTGCACCCCTGAAAGAGGTGTATGGCTTCAACCCCT 407  
 Pro Gln Leu Cys Ser Phe Ile Leu Lys Thr Ser Gln Cys Thr Leu Lys Glu Val Tyr Gly Phe Asn Pro  
 (60) (70) (80)

GAGGGCAAAGCCTTGCTGAAGAAAACCAAGAACTCTGAGGAGTTTGCTGCAGCCATGTCCAGGTATGAG 476  
 Glu Gly Lys Ala Leu Leu Lys Lys Thr Lys Asn Ser Glu Glu Phe Ala Ala Ala MET Ser Arg Tyr Glu  
 (90) (100)

CTCAAGCTGGCCATTCCAGAAGGGAAGCAGGTTTTCTTTACCCGGAGAAAGATGAACCTACTTACATC 545  
 Leu Lys Leu Ala Ile Pro Glu Gly Lys Gln Val Phe Leu Tyr Pro Glu Lys Asp Glu Pro Thr Tyr Ile  
 (110) (120)

CTGAACATCAAGAGGGGCATCATTTCTGCCCTCCTGGTTCCCCCAGAGACAGAAGAAGCCAAGCAAGTG 614  
 Leu Asn Ile Lys Arg Gly Ile Ile Ser Ala Leu Leu Val Pro Pro Glu Thr Glu Glu Ala Lys Gln Val  
 (130) (140)

TTGTTTCTGGATACCGTGTATGGAACTGCTCCACTCAGTTTACCGTCAAGACGAGGAAGGGCAATGTG 683  
 Leu Phe Leu Asp Thr Val Tyr Gly Asn Cys Ser Thr His Phe Thr Val Lys Thr Arg Lys Gly Asn Val  
 (150) (160) (170)

GCAACAGAAATATCCACTGAAAGAGACCTGGGGCAGTGTGATCGCTTCAAGCCCATCCGCACAGGCATC 752  
 Ala Thr Glu Ile Ser Thr Glu Arg Asp Leu Gly Gln Cys Asp Arg Phe Lys Pro Ile Arg Thr Gly Ile  
 (180) (190)

AGCCCCTTGCTCTCATCAAAGGCATGACCCGCCCCTGTCAACTCTGATCAGCAGCAGCCAGTCTCTGT 821  
 Ser Pro Leu Ala Leu Ile Lys Gly MET Thr Arg Pro Leu Ser Thr Leu Ile Ser Ser Ser Gln Ser Cys  
 (200) (210)

CAGTACACACTGGACGCTAAGAGGAAGCATGTGGCAGAACCATCTGCAAGGAGCAACACCTCTTCCTG 890  
 Gln Tyr Thr Leu Asp Ala Lys Arg Lys His Val Ala Glu Ala Ile Cys Lys Glu Gln His Leu Phe Leu  
 (220) (230) (240)

CCTTTCTCCTACAAGAATAAGTATGGGATGGTAGACAAGTGACACAGACTTTGAAACTTGAAGACACA 959  
 Pro Phe Ser Tyr Lys Asn Lys Tyr Gly MET Val Ala Gln Val Thr Gln Thr Leu Lys Leu Glu Asp Thr  
 (250) (260)

FIG. 2

4/9

[.....]

.....] GGA GCC GCC CCA CCG CAG CTG GCG ATG GAC CCG CCG AGG CCC GCG CTG 83 110  
MET Asp Pro Pro Arg Pro Ala Leu

CTG GCG CTG CTG GCG CTG CCT GCG CTG CTG CTG CTG CTG CTG GCG GGC GCC AGG 137 164  
Leu Ala Leu Leu Ala Leu Pro Ala Leu Leu Leu Leu Leu Leu Ala Gly Ala Arg  
signal peptide

GCC GAA GAG GAA ATG CTG GAA AAT GTC AGC CTG GTC TGT CCA AAA GAT GCG ACC 191 218  
Ala, Glu Glu Glu MET Leu Glu Asn Val Ser Leu Val Cys Pro Lys Asp Ala Thr  
↑ amino acid #1 of mature apo B protein

CGA TTC AAG CAC CTC CGG AAG TAC ACA TAC AAC TAT GAG GCT GAG AGT TCC AGT 245 272  
Arg Phe Lys His Leu Arg Lys Tyr Thr Tyr Asn Tyr Glu Ala Glu Ser Ser Ser

GGA GTC CCT GGG ACT GCT GAT TCA AGA AGT GCC ACC AGG ATC AAC TGC AAG GTT 299 326  
Gly Val Pro Gly Thr Ala Asp Ser Arg Ser Ala Thr Arg Ile Asn Cys Lys Val

GAG CTG GAG GTT CCC CAG CTC TGC AGC TTC ATC CTG AAG ACC AGC CAG TGC ATC 353 380  
Glu Leu Glu Val Pro Gln Leu Cys Ser Phe Ile Leu Lys Thr Ser Gln Cys Ile

CTG AAA GAG GTG TAT GGC TTC AAC CCT GAG GGC AAA GCC TTG CTG AAG AAA ACC 407 434  
Leu Lys Glu Val Tyr Gly Phe Asn Pro Glu Gly Lys Ala Leu Leu Lys Lys Thr

AAG AAC TCT GAG GAG TTT GCT GCA GCC ATG TCC AGG TAT GAG CTC AAG CTG GCC 461 488  
Lys Asn Ser Glu Glu Phe Ala Ala Ala MET Ser Arg Tyr Glu Leu Lys Leu Ala

ATT CCA GAA GGG AAG CAG GTT TTC CTT TAC CCG GAG AAA GAT GAA CCT ACT TAC 515 542  
Ile Pro Glu Gly Lys Gln Val Phe Leu Tyr Pro Glu Lys Asp Glu Pro Thr Tyr

ATC CTG AAC ATC AAG AGG GGC ATC ATT TCT GCC CTC CTG GTT CCC CCA GAG ACA 569 596  
Ile Leu Asn Ile Lys Arg Gly Ile Ile Ser Ala Leu Leu Val Pro Pro Glu Thr

GAA GAA GCC AAG CAA GTG TTG TTT CTG GAT ACC GTG TAT GGA AAC TGC TCC ACT 623 650  
Glu Glu Ala Lys Gln Val Leu Phe Leu Asp Thr Val Tyr Gly Asn Cys Ser Thr

CAC TTT ACC GTC AAG ACG AGG AAG GGC AAT GTG GCA ACA GAA ATA TCC ACT GAA 677 704  
His Phe Thr Val Lys Thr Arg Lys Gly Asn Val Ala Thr Glu Ile Ser Thr Glu

FIG. 3

731 758  
AGA GAC CTG GGG CAG TGT GAT CGC TTC AAG CCC ATC CGC ACA GGC ATC AGC CCA  
Arg Asp Leu Gly Gln Cys Asp Arg Phe Lys Pro Ile Arg Thr Gly Ile Ser Pro

785 812  
CTT GCT CTC ATC AAA GGC ATG ACC CGC CCC TTG TCA ACT CTG ATC AGC AGC AGC  
Leu Ala Leu Ile Lys Gly MET Thr Arg Pro Leu Ser Thr Leu Ile Ser Ser Ser

839 866  
CAG TCC TGT CAG TAC ACA CTG GAC GCT AAG AGG AAG CAT GTG GCA GAA GCC ATC  
Gln Ser Cys Gln Tyr Thr Leu Asp Ala Lys Arg Lys His Val Ala Glu Ala Ile

893 920  
TGC AAG GAG CAA CAC CTC TTC CTG CCT TTC TCC TAC AAG AAT AAG TAT GGG ATG  
Cys Lys Glu Gln His Leu Phe Leu Pro Phe Ser Tyr Lys Asn Lys Tyr Gly MET

947 974  
GTA GCA CAA GTG ACA CAG ACT TTG AAA CTT GAA GAC ACA CCA AAG ATC AAC AGC  
Val Ala Gln Val Thr Gln Thr Leu Lys Leu Glu Asp Thr Pro Lys Ile Asn Ser

1001 1028  
CGC TTC TTT GGT GAA GGT ACT AAG AAG ATG GGC CTC GCA TTT GAG AGC ACC AAA  
Arg Phe Phe Gly Glu Gly Thr Lys Lys MET Gly Leu Ala Phe Glu Ser Thr Lys

1055 1082  
TCC ACA TCA CCT CCA AAG CAG GCC GAA GCT GTT TTG AAG ACT CTC CAG GAA CTG  
Ser Thr Ser Pro Pro Lys Gln Ala Glu Ala Val Leu Lys Thr Leu Gln Glu Leu

1109 1136  
AAA AAA CTA ACC ATC TCT GAG CAA AAT ATC CAG AGA GCT AAT CTC TTC AAT AAG  
Lys Lys Leu Thr Ile Ser Glu Gln Asn Ile Gln Arg Ala Asn Leu Phe Asn Lys

1163 1190  
CTG GTT ACT GAG CTG AGA GGC CTC AGT GAT GAA GCA GTC ACA TCT CTC TTG CCA  
Leu Val Thr Glu Leu Arg Gly Leu Ser Asp Glu Ala Val Thr Ser Leu Leu Pro

1217 1244  
CAG CTG ATT GAG GTG TCC AGC CCC ATC ACT TTA CAA GCC TTG GTT CAG TGT GGA  
Gln Leu Ile Glu Val Ser Ser Pro Ile Thr Leu Gln Ala Leu Val Gln Cys Gly

1271 1298  
CAG CCT CAG TGC TCC ACT CAC ATC CTC CAG TGG CTG AAA CGT GTG CAT GCC AAC  
Gln Pro Gln Cys Ser Thr His Ile Leu Gln Trp Leu Lys Arg Val His Ala Asn

1325 1352  
CCC CTT CTG ATA GAT GTG GTC ACC TAC CTG GTG GCC CTG ATC CCC GAG CCC TCA  
Pro Leu Leu Ile Asp Val Val Thr Tyr Leu Val Ala Leu Ile Pro Glu Pro Ser

1379 1406  
GCA CAG CAG CTG CGA GAG ATC TTC AAC ATG GCG AGG GAT CAG CGC AGC CGA GCC  
Ala Gln Gln Leu Arg Glu Ile Phe Asn MET Ala Arg Asp Gln Arg Ser Arg Ala

ACC TTG TAT GCG CTG AGC CAC  
Thr Leu Tyr Ala Leu Ser His

# Sal-Sma-Linker Region

Arg, Arg, Gln, Lys, Gln, Leu, Asn  
 R, R, Q, K, Q, L, N  
 GTCCGACAAACAGCTAAAGTAAGGACCCAGCCTGGGGTTGAGGGC  
 Sa/I TTTTGTGCAATTTATTCCTGGGTCGGACCCCAACTCCCG

SmaI

AGGGGCGAGGGGGCAGGG  
 TCCCGTCCCGTCCC

oligo's:	GTCGACAAACAGCTAAAGTAAGGACC	28 mer	# 1
	CAGCCCTGGGGTTGAGGGC	18 mer	# 2
	AGGGGCGAGGGGGCAGGG	17 mer	# 3
	CCCTGCCCCCTGCCCTCA	24 mer	# 4
	ACCCCAAGGCTGGGTCCCT	17 mer	# 5
	TACTTTAGCTGTTTTT	16 mer	# 6

FIG. 4A

map of pPSAI(2)

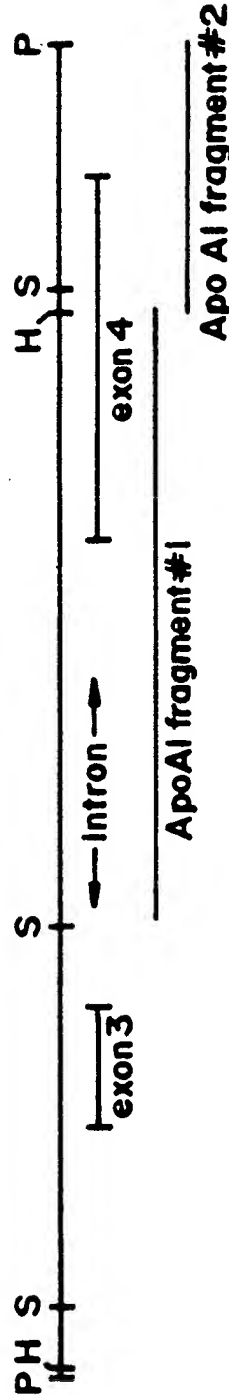


FIG. 4B

29 56  
 GCC GCT GAG GAG CCC GCC CAG CCA GCC AGG GCC GCG AGG CCG AGG CCA GGC CGC

83 110  
 AGC CCA GGA GCC GCC CCA CCG CAG CTG GCG ATG GAC CCG CCG AGG CCC GCG CTG  
 MET Asp Pro Pro Arg Pro Ala Leu

137 164  
 CTG GCG CTG CTG GCG CTG CCT GCG CTG CTG CTG CTG CTG GCG GGC GCC AGG  
 Leu Ala Leu Leu Ala Leu Pro Ala Leu Leu Leu Leu Leu Ala Gly Ala Arg  
 signal peptide

191 218  
 GCC GAA GAG GAA ATG CTG GAA AAT GTC AGC CTG GTC TGT CCA AAA GAT GCG ACC  
 Ala, Glu Glu Glu MET Leu Glu Asn Val Ser Leu Val Cys Pro Lys Asp Ala Thr  
 amino terminus of mature protein

245 272  
 CGA TTC AAG CAC CTC CGG AAG TAC ACA TAC AAC TAT GAG GCT GAG AGT TCC AGT  
 Arg Phe Lys His Leu Arg Lys Tyr Thr Tyr Asn Tyr Glu Ala Glu Ser Ser Ser

299 326  
 GGA GTC CCT GGG ACT GCT GAT TCA AGA AGT GCC ACC AGG ATC AAC TGC AAG GTT  
 Gly Val Pro Gly Thr Ala Asp Ser Arg Ser Ala Thr Arg Ile Asn Cys Lys Val

353 380  
 GAG CTG GAG GTT CCC CAG CTC TGC AGC TTC ATC CTG AAG ACC AGC CAG TGC ATC  
 Glu Leu Glu Val Pro Gln Leu Cys Ser Phe Ile Leu Lys Thr Ser Gln Cys Ile

407 434  
 CTG AAA GAG GTG TAT GGC TTC AAC CCT GAG GGC AAA GCC TTG CTG AAG AAA ACC  
 Leu Lys Glu Val Tyr Gly Phe Asn Pro Glu Gly Lys Ala Leu Leu Lys Lys Thr

461 488  
 AAG AAC TCT GAG GAG TTT GCT GCA GCC ATG TCC AGG TAT GAG CTC AAG CTG GCC  
 Lys Asn Ser Glu Glu Phe Ala Ala Ala MET Ser Arg Tyr Glu Leu Lys Leu Ala

515 542  
 ATT CCA GAA GGG AAG CAG GTT TTC CTT TAC CCG GAG AAA GAT GAA CCT ACT TAC  
 Ile Pro Glu Gly Lys Gln Val Phe Leu Tyr Pro Glu Lys Asp Glu Pro Thr Tyr

569 596  
 ATC CTG AAC ATC AAG AGG GGC ATC ATT TCT GCC CTC CTG GTT CCC CCA GAG ACA  
 Ile Leu Asn Ile Lys Arg Gly Ile Ile Ser Ala Leu Leu Val Pro Pro Glu Thr

623 650  
 GAA GAA GCC AAG CAA GTG TTG TTT CTG GAT ACC GTG TAT GGA AAC TGC TCC ACT  
 Glu Glu Ala Lys Gln Val Leu Phe Leu Asp Thr Val Tyr Gly Asn Cys Ser Thr

677 704  
 CAC TTT ACC GTC AAG ACG AGG AAG GGC AAT GTG GCA ACA GAA ATA TCC ACT GAA  
 His Phe Thr Val Lys Thr Arg Lys Gly Asn Val Ala Thr Glu Ile Ser Thr Glu

FIG. 5

731 758  
AGA GAC CTG GGG CAG TGT GAT CGC TTC AAG CCC ATC CGC ACA GGC ATC AGC CCA  
Arg Asp Leu Gly Gln Cys Asp Arg Phe Lys Pro Ile Arg Thr Gly Ile Ser Pro

785 812  
CTT GCT CTC ATC AAA GGC ATG ACC CGC CCC TTG TCA ACT CTG ATC AGC AGC AGC  
Leu Ala Leu Ile Lys Gly MET Thr Arg Pro Leu Ser Thr Leu Ile Ser Ser Ser

839 866  
CAG TCC TGT CAG TAC ACA CTG GAC GCT AAG AGG AAG CAT GTG GCA GAA GCC ATC  
Gln Ser Cys Gln Tyr Thr Leu Asp Ala Lys Arg Lys His Val Ala Glu Ala Ile

893 920  
TGC AAG GAG CAA CAC CTC TTC CTG CCT TTC TCC TAC AAG AAT AAG TAT GGG ATG  
Cys Lys Glu Gln His Leu Phe Leu Pro Phe Ser Tyr Lys Asn Lys Tyr Gly MET

947 974  
GTA GCA CAA GTG ACA CAG ACT TTG AAA CTT GAA GAC ACA CCA AAG ATC AAC AGC  
Val Ala Gln Val Thr Gln Thr Leu Lys Leu Glu Asp Thr Pro Lys Ile Asn Ser

1001 1028  
CGC TTC TTT GGT GAA GGT ACT AAG AAG ATG GGC CTC GCA TTT GAG AGC ACC AAA  
Arg Phe Phe Gly Glu Gly Thr Lys Lys MET Gly Leu Ala Phe Glu Ser Thr Lys

1055 1082  
TCC ACA TCA CCT CCA AAG CAG GCC GAA GCT GTT TTG AAG ACT CTC CAG GAA CTG  
Ser Thr Ser Pro Pro Lys Gln Ala Glu Ala Val Leu Lys Thr Leu Gln Glu Leu

1109 1136  
AAA AAA CTA ACC ATC TCT GAG CAA AAT ATC CAG AGA GCT AAT CTC TTC AAT AAG  
Lys Lys Leu Thr Ile Ser Glu Gln Asn Ile Gln Arg Ala Asn Leu Phe Asn Lys

1163 1190  
CTG GTT ACT GAG CTG AGA GGC CTC AGT GAT GAA GCA GTC ACA TCT CTC TTG CCA  
Leu Val Thr Glu Leu Arg Gly Leu Ser Asp Glu Ala Val Thr Ser Leu Leu Pro

1217 1244  
CAG CTG ATT GAG GTG TCC AGC CCC ATC ACT TTA CAA GCC TTG GTT CAG TGT GGA  
Gln Leu Ile Glu Val Ser Ser Pro Ile Thr Leu Gln Ala Leu Val Gln Cys Gly

1271 1298  
CAG CCT CAG TGC TCC ACT CAC ATC CTC CAG TGG CTG AAA CGT GTG CAT GCC AAC  
Gln Pro Gln Cys Ser Thr His Ile Leu Gln Trp Leu Lys Arg Val His Ala Asn

1325 1352  
CCC CTT CTG ATA GAT GTG GTC ACC TAC CTG GTG GCC CTG ATC CCC GAG CCC TCA  
Pro Leu Leu Ile Asp Val Val Thr Tyr Leu Val Ala Leu Ile Pro Glu Pro Ser

GCA CAG CAG CTG CGA GAG ATC] CGT CGA CAA AAA CAG CTA.....  
Ala Gln Gln Leu Arg Glu Ile] Arg Arg Gln Lys Gln Leu.....



FIG. 5  
(con't)

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US86/02074

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (if several classification symbols apply, indicate all) <sup>3</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC (4) C12P 21/00 U.S. 435/68		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>4</sup>		
Classification System	Classification Symbols	
U.S.	435/68, 172.3, 253, 320 935/10, 11, 15, 47 424/38, 85, 88 530/359, 389, 391, 402, 403, 405	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched <sup>5</sup>		
BIOTECHNOLOGY RESEARCH PARTNERS, LTD.		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <sup>14</sup>		
Category <sup>6</sup>	Citation of Document, <sup>15</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
X Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, Volume 75 (10), October 1978, Krieger et al, "Receptor-Mediated Uptake of Low Density Lipoprotein Reconstituted with 25-Hydroxycholestearyl Oleate Suppresses 3-Hydroxy-3-Methylglutaryl-Co-enzyme A Reductase and Inhibits Growth of Human Fibroblasts", pages 5052-6.	<u>1-7&amp;13</u> 8-12
X Y	CANCER RESEARCH, Volume 43, October 1978, Rudling et al, "Delivery of Aclacinomycin A to Human Glioma Cells in Vitro by the Low-Density Lipoprotein Pathway", pages 4600-5.	<u>1-7&amp;13</u> 8-12
Y	WO, B, 8000030 (HARVARD COLLEGE) 10 January 1980, see page 2 line 5 -page 3 line 24.	8-12
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p><sup>6</sup> Special categories of cited documents: <sup>15</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 45%;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"&amp;" document member of the same patent family</p> </div> </div>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>1</sup> 16 December 1986		Date of Mailing of this International Search Report <sup>2</sup> <b>02 JAN 1987</b>
International Searching Authority <sup>1</sup> ISA/US		Signature of Authorized Officer <sup>20</sup> John E. Tarcza <i>John E. Tarcza</i>

## III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No <sup>18</sup>
X Y	JOURNAL OF MEDICINAL CHEMISTRY, Volume 27, 1984, Firestone et al, "Selective Delivery of Cytotoxic Compounds to Cells by the LDL Pathway", pshrd 1037-43.	<u>1-7&amp;13</u> 8-12
X Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, Volume 82, July 1985, Lusi et al, "Cloning and Expression of Apolipo- protein B, the Major Protein of Low and Very Low Density Lipoproteins", pages 4597-4601.	<u>13</u> 8-12
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, Volume 77 (10), October 1980, Dolby et al, "Cloning and Partial Nucleotide Sequence of Human Immunoglo- bulin Mu Chain cDNA from B Cells and Mouse-Human Hybridomas" pages 6027-31.	8-13
A	US, A, 4,275,000 (ROSS) 23 June 1981, See column 8 lines 6-24 and example 3.	1-7
A	US, A, 4,429,008 (MARTIN) 31 January 1984, See example VI.	1-7